黄姓及相关姓氏族谱重构：分子遗传学与历史文献的综合考察

黄石

中南大学生命科学学院医学遗传学研究中心

关键词：黄姓，徐姓，梁姓，伯益，嬴姓，黄帝，夏朝，Y 染色体单倍型，M117-F8-F2137-A16635

摘要

黄姓的源流比较复杂，但根据《史记》等史书的记载，黄姓可追溯至黄帝后裔伯益，伯益因协助大禹治水被赐嬴姓。他有二子，长子大廉建立了黄国，位于河南潢川，次子若木建立了徐国，位于江苏宿迁。黄国被灭后，国民以国名为姓，并迁徙至湖北地区；徐国被灭后，国民以徐为姓，散居江苏；伯益更远的子孙还衍生出了包括梁国梁姓（灭国后散居晋国今山西河北）在内的其他 12 个嬴姓下属姓氏。但鉴于上古时期的文献叙事往往被视为传说性质，本研究致力于通过分析黄姓及其相关姓氏的父系基因单倍型，以科学的方法独立验证文献记载的客观性。我们分析了 23 魔方公司公布的 207333 汉族男性的姓氏基因数据，并结合了独立测序研究。在每个有黄姓、徐姓、或梁姓至少 5 人的、形成时间为 4000 年前左右的共 189 个 Y 染色体单倍型支系，我们分析了该支系湖北黄姓在全国同支系黄姓人数的占比，占比高的支系就可能是有湖北代表性的黄姓支系。我们也同样分析了苏沪徐姓在全国同支系徐姓人数的占比，以及山西河北梁姓在全国同支系梁姓人数的占比。我们进一步采用了两个标准来筛选某地域特异的某姓氏始祖基因型：一是该姓氏应在该基因型的所有姓氏中有高频出现；二是该地域内的样本在该基因型应为占比较高。结果显示，黄姓在单倍型 O-M117-F8-F2137-A16635-A16636 下游的 MF14296 或 MF15137（共祖时间 3420-4320 年前）显示了湖北高频特征。我们独立采样研究发现，在武汉黄陂罗汉寺周边的九个同族黄姓村庄中，A16636 是黄姓族人的主流基因。在全国范围内，MF14296/MF15137 支系中黄姓占比较高，湖北籍样本也占比较高。另外，我们发现，属于 A16636 姐妹支系的 F15823 下游的 MF38096/SK1726 支系（共祖时间 2830-5200 年前）的徐姓人士都来自苏沪，属于 F15823 下游的 MF14287/MF15398/MF14963 支系（共祖 1850-4900 年前）的梁姓人士基本都来自河北。基于上述结果，黄姓始祖可能是黄国某个带有 MF14296/MF15137 的黄姓公族家族，徐姓始祖可能是徐国某个带有 MF38096/SK1726 的徐姓公族家族，梁姓始祖可能是梁国某个带有 MF14287/MF15398/MF14963 的梁姓公族家族。这一结论与《史记》等文献的记载较为一致，为这些有关黄国、徐国、和夏朝叙事的客观性提供了有力的遗传学佐证。

前言

中国汉人的姓氏大概 5000 年前时开始形成 (1-3)。4000 年前时存在有上古八大姓，他们被认为是今天 ~23000 个姓氏的祖先 (4)。其中的嬴姓就被认为是今天 14 个姓氏的祖先，包括黄、徐、梁、赵、秦、江、葛、翟、谷、马、肖、钟、缪、费。这其中有 5 个今天中国人口排名前 22 的大姓，包括黄 （排名第 7），赵（第 8），徐（第 11），马（第 13），梁（第 22）。汉朝历史学家司马迁（前 145 年—前 86 年？），可以说是中国历史上最伟大的历史学家，在他的伟大巨著《史记》中关于秦国的起源有这样的记载："秦之先为嬴姓。其后分封，以国为姓，有徐氏、郯氏、莒氏、终黎氏、运奄氏、菟裘氏、将梁氏、黄氏、江氏、修鱼氏、白冥氏、蜚廉氏、秦氏。然秦以其先造父封赵城，为赵氏 "。这显示，在这 14 个被史记记载的嬴姓下属姓氏中，与今天姓氏存在明确对应关系的古姓氏包括徐、将梁（即梁）、黄、江、秦、和赵等 6 个姓。

依据《古本竹书纪年》等史书记载，黄氏发源于中国内蒙古东部、燕山之南的辽河西源－西拉木伦河（原名潢水）流域，起源于栖息该河沿岸的一个部落黄夷族，属于黄帝长子少昊部落，后黄夷族向东南

迁徙，经河北的黄山和黄丘进入山东半岛，加入东夷集团，成为九夷中的一支。据《史记》记载，帝舜时代，少昊后人、东夷部落的首领伯益，因帮助夏朝开国君王大禹治水有功，被帝舜赐姓嬴。伯益有二子，公元前 21 世纪夏代初年（亦说公元前 2148 年），长子大廉于光州定城西十二里隆古乡建立黄国（今河南潢川），世袭相传至公元前 648 年为楚所灭，亡国后的黄国民众以国为姓。南宋郑樵所著的《通志・氏族略》关于黄国的记载如下："黄氏，嬴姓，少昊之后也。少昊金天氏，其后有黄国，春秋时为楚所灭，子孙以国为氏。"

至于黄国为何名黄，可能与伯益来自远古黄夷族部落有关，该部落喜欢用"黄"来命名生存之地，例如，在山东、河南、湖北等地，就有许多以"黄"为名的地名，如山东黄县，河南潢川县，湖北黄陂区等等，这些地名可能与黄夷族部落的迁徙和居住历史有关。 另外，可能与大廉是长子有关，长子往往承载着家族的传统和期望，这种传统可能在中国历史上很早以前就开始了。

伯益次子徐若木建立了徐国（今江苏宿迁），世袭相传至公元前 512 年国家为吴国所灭。《通志・氏族略》关于若木的记载如下："伯益佐禹有功，封其子若木于徐，子孙以国为氏。"

嬴姓十四氏只有黄姓和徐姓是直接来自伯益的儿子，其它姓氏则是来自与伯益较远的子孙，比如伯益 14 世孙造父被记载是嬴姓赵氏和嬴姓马氏的始祖。另外，根据《史记》和《通志・氏族略》等史书记载，伯益的后裔有个叫非子的人 （公元前 858 年去世），周孝王让他负责养马，封秦邑给他，被称为秦嬴。非子的曾孙秦仲的小儿子被封在梁，建立了梁国（今天陕西韩城），称梁康伯。后来秦国灭了梁国。

黄国被楚国灭后黄姓族群随之扩散，主要流向附近的湖北。在湖北东部，自西向东分别排列有城市黄陂、黄安（今红安）、黄冈、黄石、和黄梅。其中，黄陂和黄冈与黄国关系紧密(5)，黄陂曾是黄国的所辖区域，距离河南潢川黄国遗址较近。黄冈取自黄冈山，据《今县释名》："西北有黄冈山，因古黄国得名。"因此，探讨湖北地区的黄姓源流对于探讨黄姓源流和黄国乃至夏朝历史具有重要意义。

《通志・氏族略》提到："自若木至偃王三十二世，为周所灭，复封其子宗为徐，子宗十一世章羽，昭三十年，为吴所灭，子孙以国为氏。" 这段记载了徐国从夏朝的若木一直传了 32 世代后到周朝的徐偃王，徐国最终被吴国灭亡后，国民以徐为姓。徐姓族群基本留在了江苏，苏沪地区是今天徐姓人口最高频的地区。因此，探讨苏沪地区的徐姓源流对于探讨徐姓源流和徐国乃至夏朝历史具有重要意义。另外，研究徐姓源流也可以为黄姓源流研究的结论提供独立的旁证。

亡国后的梁国子孙，大部分逃到了晋国，国名为姓。晋国当时包含山西和河北。因此，探讨山西河北地区的梁姓历史对于认识梁姓源流和梁国历史具有重要意义，也可以为黄姓和徐姓源流提供旁证。

《史记》等史书中关于黄姓、黄国、徐姓、徐国、伯益以及夏朝的记载，如果其所述内容属实，则理论上应当能够通过不同领域独立的研究进行验证。分子遗传学作为一种科学手段，能够追溯姓氏的始祖及其大致的形成年代，因此它提供了一种独立验证历史文献中关于某姓氏源流记载的可能性。本研究的主要目的是运用姓氏遗传学的方法来论证黄姓及相关姓氏的源流，并探讨分子遗传学结论与《史记》等历史文献记载之间的一致性。

方法与材料:

23 魔方数据分析

基因公司 23 魔方公开了 22 万人的父系单倍型数据（数据截止日期 2024 年 11 月 8 日），其中有汉人 207333 人，有部分包括了高通量测序。我们通过魔方网站的筛选功能（姓氏、省市等）系统分析了这些数据。

在每个有魔方客户黄姓、徐姓、或梁姓至少 5 人的、形成时间为 4000 年前左右的主要支系，我们分析了该支系湖北黄姓人数在全国该支系黄姓人数的占比，占比高的支系就有可能是湖北最普遍的或最有湖北代表性的黄姓支系，也就较有可能是最初源自黄国并流向湖北的黄姓支系，也就应该是最符合文献记载的黄姓源流的支系。同样，我们也分析了该支系江苏上海（苏沪）徐姓人数在全国同支系徐姓人数的占比，占比高的支系就有可能是最有苏沪代表性的徐姓支系。我们还分析了该支系山西或河北的梁姓人数在全国同支系梁姓人数的占比，占比高的支系就有可能是最有山西或河北代表性的梁姓支系。这是从分子遗传学独立无偏向地探讨黄姓、徐姓、和梁姓源流，若结果不符合记载，则意味着对记载的挑战，若符合，则支持记载的可靠和真实性。

对于以上找到的支系，我们进一步分析了该支系的黄姓、徐姓、或梁姓是否在同支系全国所有姓氏的占比有较高比例，只有占比较高的被视为是可能性最大的某姓氏的始祖支系。同时，我们也分析了这些支系的所属省份的人数在全国的占比，占比较高的支系被认为是最有可能是始祖的支系。只有同时满足这三个条件的支系才被认为是某省份最具地域代表性的、最有可能包含有姓氏始祖的支系。

作为湖北特异黄姓支系鉴定的对照，我们也同样分析了是否存在类似的江苏黄姓特异的支系。

客户中如果出现 2 个或 2 个以上的同一姓氏、同一城区、同一账号的情况，则被认为是相关者属于亲属关系，仅按一人计算，并针对关键数据与魔方公司相关职员进行了核实。

另外，大部分魔方客户没有高通测序或没有 A16636 下游的信息，为了研究需要， 我们赞助了 7 位相关客户的高通测序，包括 A16636 阳性的黄姓 6 位中的 3 位。

样本收集与 DNA 提取和测序

我们采集 7 个祖籍武汉黄陂的黄姓人士的唾液样本，分属三个不同村庄，唾液 DNA 提取试剂盒购自 Norgen Biotek 公司。DNA 提取和 PCR 扩增以及 PCR 片段测序都是参照常见方法。测序位点为 A16636 等位基因，位于基因组参考联盟人类基因组构建 37 版（GRCh37）Y 染色体序列位置 8187115， PCR 上游引物为 5’－TGACACATTGCTGAGCCCAA-3’ ，下游引物为 5’－ TGTGAGCCAAAACACCTACCA-3’。

我们还采集了一个祖籍武汉黄陂的黄姓人士的血液样本进行了全基因组深度测序，由诺禾致源公司完成，并把 Y 染色体序列数据上传给 23 魔方公司整合进 23 魔方数据库。

相关人士都签署了知情同意书。

结果:

湖北黄姓特异父系基因支系的鉴定

基因测序公司 23 魔方公开了 207333 名汉族男性的父系单倍型数据（数据截止日期 2024-11-18），其中黄姓汉人占比 0.024（即 4957 人），这一比例接近 2010 年全国第六次人口普查中黄姓人口在全国的占比 0.022（表一）。另外，广东和湖北黄姓用户人数在全国黄姓的占比也基本符合人口普查的结果。具体来说，广东黄姓的占比 0.234 与人口普查中黄姓在广东的占比 0.195 相近，湖北黄姓的占比 0.047 也与人口普查中相应的比例 0.059 接近。因此，魔方客户中黄姓的分布情况基本上符合人口普查结论的预期，没有出现黄姓人口的特别富集或贫乏现象。同时，也没有发现黄姓人口较多的广东省或湖北省的黄姓人数占比过度高于或低于正常预期的情况。作为对照比较，全国陈姓客户在全部用户中占比 0.05 也符合人口普查陈姓在总人口占比 0.0453，湖北陈姓客户在全国陈姓的占比 0.039 也接近人口普查湖北陈姓在全国陈姓的占比 0.0514，广东陈姓客户占在全国陈姓的占比 0.154 也接近人口普查比例 0.174。

湖北黄姓特异的父系基因支系应该具有湖北黄姓高频的特征，即湖北黄姓在该支系的占比高于其它地区黄姓在该支系的占比。在每个有魔方客户黄姓至少 5 人的、形成时间为 4000 年前左右的主要支系，我们分析了该支系湖北黄姓人数在全国该支系黄姓人数的占比，占比高的支系就有可能是湖北最普遍的或最有湖北代表性的黄姓支系，也就较有可能是最初源自黄国并流向湖北的黄姓支系，也就应该是最符合文献记载的黄姓源流的支系。

我们分析了约占汉族 75%的主流单倍型 O-F175 下游各个支系的黄姓分布情况，这些支系包括了三大超级祖先基因型 M117、F46 及 F11 的下游支系 (6)（表一,表二，和附件材料表一）， 我们也分析了其它所有的单倍型包括欧洲和非洲人的典型单倍型（表二，附件表一）。作为黄姓的对照，我们也分析了人口数稍高于黄姓的南方大姓陈姓的分布（表一）。对于绝大部分单倍型支系，黄姓的支系占比均接近人口预期，陈姓占比也较符合预期。例如，黄姓在 M117 占比为 0.025，接近魔方客户的人口比例 0.024。

黄姓在东亚较普遍的各类单倍型均有出现，包括 O，N，Q，C，和 D，也在部分欧洲和非洲较普遍的单倍型中出现，如 R 和 E，但在大部分非东亚典型单倍型中没有出现，包括 A、B、G、H、I、J、L、M、S、和 T （附件表一）。黄姓这种在各个单倍型的广泛的分布是大姓的一个普遍现象，说明大部分黄姓并不共享一个较近的始祖，或一个 4000 年前左右姓氏开始形成时期的始祖。这其实符合《史记》关于黄姓起源的记载，一个国家的国民自然会包括各种人或各种遗传背景不同的人，若这些人在国家灭亡时都以国名为姓，则黄国的以黄为姓的国民的遗传背景或父系单倍型自然应该是非常多样的。但其中应该有一个主流的或较常见的黄姓代表单倍型，这个单倍型的人口较多，属于较大的家族，这样的家族最有可能是世袭多代的国君家族。如果黄国国民和国君家族在国家灭亡后都大部分迁徙到了湖北，则其中较普遍的或公族家族的黄姓所属单倍型应该具备湖北特异性高频的特征。

结果显示，在 A16636 这个支系，湖北黄姓在全国同姓占比为 0.5，较显著高于预期占比 0.047 （表一）。另外，距离 A16636 较近的上游祖先单倍型（A16635， F2137，F2643，F1754，和 ACT7481）也显示了湖北黄姓占比较高于预期的现象（0.14 左右）。作为湖北黄姓的对照，我们也分析了广东黄姓在表一所示的 M117 下游各个主要支系的分布，但没有发现存在有广东黄姓在全国黄姓占比较高的支系。

在 A16636 下游的 MF14296 支系也显示了湖北黄姓占比高（0.5），并且黄姓在这个支系的所有姓氏中的占比也比较高，为 0.11（表二），较显著高于全国人口比例 0.022。还有，MF14296 支系的湖北样本占比为 0.2 （表二），较显著高于人口普查结果显示的湖北占比 0.043，显示该支系具备湖北特异性高频的特征。这些结果提示，这只湖北特异的黄姓基因型 A16636-MF14296 符合史书关于黄姓源流的记载。

如附件表一所示，在被检测的一共 189 个支系中 ，绝大部分没有展现出黄姓在湖北高频的特征，仅 C-MF1017 及其下游 C-Z45203 是例外，频率分别是 0.53 和 0.57 （表二） 。但 MF1017 支系的黄姓人数在全国所有姓氏中占比不高，仅是 0.05，未显示黄姓人口富集的特征 （表二）。此外，这个支系的湖北人

数在全国的占比不高 （0.1），并未呈现出湖北特异性高频的特征。C-Z45203 支系的黄姓人数在全国所有姓氏中占比较高（0.09）（表二），但这个支系的湖北人数在全国的占比不高（0.084）。最后，值得注意的是，这两个支系的黄姓大部分来自湖北黄石，在湖北五个黄姓城市中，黄石的地理位置属于是距离黄国故城较远的一个。因此，尽管 C-MF1017 及其下游支系 Z45203 在一定程度上表现出湖北地区的特异性特征，但相较于之前提及的 MF14296 支系，它们的湖北特异性略显不足。

作为对照，我们分析了各个支系的江苏黄姓在全国黄姓的占比 (附件表一)，但没有发现存在有江苏黄姓高频的支系(等于或高于 0.5)，虽然有一支 O-F4068 较高频（0.46），但这个支系的黄姓在所有姓氏占比仅为 0.026 （附件表一），与人口普查比例类似，这说明湖北存在湖北黄姓高频的支系具有一定的省市特异性。

在分析 23 魔方数据库之前，我们曾通过高通测序分析了某湖北黄姓个体，并上传到了魔方数据库，其最末端遗传标记是 A16636 下游的 MF287231 （图一）。他的祖父家在武汉黄陂区罗汉寺黄家墩村，村里大姓是黄，根据族谱记载，同族黄姓宗亲还包括罗汉寺区域的其它 8 个以黄命名的村庄，黄八塆、黄家大塆、黄家岗、黄家垱、黄家院子、黄家大独屋、黄家小独屋、及黄胡细塆(7)。为了验证这些村民的主流基因是否为 A16636，我们进行了采样测序实验。我们随机采集了 7 个黄姓样本，包括黄家墩村三人，黄家大湾村两人，以及无族谱关系的、村名中不带有黄字的邻近村庄两路口村两人。结果显示，黄家墩村 3 人中有 2 人是 A16636，一人不是，但该人经后续调查得知是属于领养。黄家大湾 2 人都是 A16636，两路口村 2 人均不是 A16636。这些结果说明，黄家墩村所代表的九个村的黄姓同族宗亲应该都是 A16636 基因型。

魔方数据中 A16636 全国有 6 个黄姓 （数据库显示 7 人，但其中两个宁夏黄是父子关系，故仅按一人算），其中 3 个没有下游支系测序信息，被我们赞助后完成了深度测序（宁夏黄姓父子两人中仅测了父亲）。结果显示，所有 6 个个体都是属于支系 MF14296。与 MF287231 个体最近的两个支系是某武汉汉阳区的黄姓个体 MF339314 和某湖北随州黄姓个体 MF592119（汉阳位于黄陂南部 40 公里，随州位于黄陂西北 80 公里），三者的最近共同祖先单倍型为 MF315409，共祖时间推算是 510-1210 年前。这两个支系家族不属于是黄家墩村所属族人的已知亲属关系人。这个结果说明，黄家墩村所属黄姓族人至少在明元宋时期（510-1210 年前）就有了一个居住在湖北东北部邻近河南东南部潢川黄国的黄姓祖先。

湖北黄姓支系的姓氏起始年代

湖北 A16636 下游的黄姓支系的命名祖先应该是在什么年代或哪个支系开始？鉴于有 3 个湖北黄姓都带有 MF315409，共祖时间是 510-1260 年，可以推断当年他们应该有一个居住在湖北的黄姓祖先。这个支系的上游是 MF309877（共祖时间 1260-2080 年），某湖北孝感陶姓 MF582799 也带有这个标记。这提示当年有一个 MF309877 型湖北人，其后代也居住在湖北，大部分黄姓，小部分陶姓。这个祖先大概率应该姓黄，其黄姓后裔的一小部分后来改姓了陶。与这个推论一致，A16636 湖北陶姓仅占全国同支系陶姓的 0.25，没有显示出湖北较高富集的现象（A16636 湖北黄姓则占了全国同支系黄姓的 0.5）。

与 MF309877 支系最近的黄姓支系是一个徐州黄姓支系，属于 MF436936 支系。这两个不同黄姓支系的共祖标记是 MF607191，共祖时间是 2950-3310 年，该支系下游有 41 人，11 人（0.268）来自湖北，其它省份最多是 4 人，说明该支系人口较集中在湖北，也说明其祖先应该是生活在湖北，由于其湖北后裔中黄姓占比较高（0.27），可推导该祖先应该姓黄，其湖北后裔的其它姓氏可能是来自改姓和其它因素。

MF217876 是 MF607191 最近的上游，一个北京黄姓带有 MF217876 但不带有 MF607191，MF217876 的共祖时间是 3310-3420 年，这个支系比 MF607191 支系只多出了 3 人，但其中就有黄姓 1 人，这提示带有 MF217876 的湖北后裔的湖北祖先应该姓黄，其后裔中有部分黄姓迁徙到其它地区。

MF15137 是 MF217876 最近的上游，仅比 MF217876 多出 2 人，共祖时间是 3420-3610 年，因此，带有 MF15137 的湖北后裔的湖北祖先大概率应该姓黄。

MF14296 是 MF15137 的最近上游和 A16636 的最近下游，比 MF15137 多出 10 人（共 56 人），湖北占 11 人（0.20），为全国最高，其它较高省市是江苏 5 人，山东 5 人，河北 5 人，浙江 4 人，安徽 3 人，河南 3 人。该支系的黄姓在全国同支系所有姓氏占比为 0.11，这提示带有 MF14296 的湖北后裔的湖北祖先大概率应该姓黄。其它姓氏在 MF14296 支系人数较多的其它姓氏包括张 5 人、任 5 人、李 4 人。相比之下，A16639 是 MF14296 的姊妹支，但没有湖北占比高的现象（0.018），进一步佐证 MF14296 支系的湖北特异性。

A16636 的湖北人在全国占比是 0.042，该支系黄姓在全国同支系所有姓氏占比为 0.008，小于人口预期比例 3 倍，因此，带有 A16636 的祖先的大部分后裔没有湖北集中的现象，也没有黄姓集中的现象，所以，带有 A16636 的个体大概率应该不姓黄，也不居住在湖北。

综合以上分析，A16636 下游的湖北黄姓人口的以黄为姓氏的最早祖先标记可能是 MF14296，但也不能排除是 MF15137,这两个支系共祖时间范围是 3420-4320 年前，属于黄国历史的前期或中期。因此，以黄为姓应该是在黄国灭亡之前，也就只能是黄国公族的姓氏。这就说明，黄国国君的支系可能是 MF14296 或 MF15137。

徐姓分析

如果黄姓基因型历史符合《史记》记载的黄姓出自嬴姓伯益的长子大廉，则徐姓，作为伯益次子若木的子孙，就是最有可能也具备类似符合《史记》记载的基因型历史。徐姓来自徐国国民以国名为姓，聚集在江苏。因此，徐姓始祖支系应该在江苏具有特异性高频的特征。

类似对黄姓的研究，我们分析了共 189 个单倍型支系，计算了每个支系的江苏上海（苏沪）徐性在全国同姓人口中的比例（表二，附件表一）。结果表明，苏沪徐姓富集在支系 A16635-F15823-F1442-MF38096（共祖时间 3370-5200 年前）， 该支系苏沪徐姓在全国同支系徐姓人口占比为 100%， 徐姓在所有姓氏占比 0.104，显著高于全国人口比例 0.014，该支系的苏沪人口在全国的占比也较高 （0.25）。

我们还鉴定出了 23 个其它支系，这些支系各自的苏沪徐姓在全国同支系徐姓人数的占比都比较高（大于 0.5，表二，附件表一），但从徐姓在所有姓氏的占比这个指标看，这些支系都处于较低水平（0.008-0.057），因此并不属于徐姓较显著富集的支系。虽然这些支系中有部分具有苏沪人口高频的特征，但他们都没有富集徐姓人口，例如，苏沪人口在支系 O-Y89945 占比高达 0.43，但该支系中徐姓在所有姓氏的占比仅为 0.015，因此，这些支系相较于支系 O-MF38096 而言，均不是最具苏沪代表性的徐姓支系。

MF38096 的上游支系 F1442 虽然有苏沪徐姓在全国同姓占比高的特征 （0.7），但没有徐姓在所有姓氏中的富集特征（仅为 0.018）（表二）。MF38096 的徐姓样本都集中在其下游支系 SK1726（共祖时间 2830-3480 年前）（表二）， 但这些徐姓样本都没有 SK1726 下游的信息。 SK1726 的样本数是 40 人，占 MF38096 的样本数的 73%，这两个支系都较符合徐姓始祖支系的标准，但 MF38096 的上游 F1442 就明显不符合，因此，徐姓始祖支系有可能是 MF38096，但也不能排除是 SK1726。这与史书记载的徐国和徐姓

源流基本一致。这一徐姓支系与上面鉴定的黄姓支系的共享最近始祖是 A16635，进一步印证了黄姓来自嬴姓的《史记》记载。

梁姓分析

梁姓属于是伯益较远的子孙非子的后人，梁国建国时间在约公元前 800 年，梁国被灭后，国民以梁为姓迁徙去了晋国，也即今山西河北一带。同上述分析类似，我们研究了共 189 个支系，分析了每个支系山西或河北梁姓在全国梁姓人口的占比。结果发现有 5 个支系出现了河北梁姓在全国梁姓占比高的情况（高于 0.5）(表二)，其中只有 MF14287 支系（4090-4900 年前）及其二个下游支系 MF15398 和 MF14963 ，具有梁姓在所有姓氏中占比高的情况（0.1），显著高于梁姓人口比例 0.007。但是，MF14287 的上游支系 MF15764 的梁姓在所有姓氏的占比仅为 0.015 (表二)。进一步发现，梁姓都是属于 MF14963 支系（共祖时间 1850-3210 年前），在该支系所有姓氏中占比为 0.1，另外，该支系河北人口占比最高（0.21），该支系下游至少有两个不同支系含有梁姓样本。这些结果提示，梁国梁姓的始祖支系比较可能是 MF14963，但也不能排除是 MF14287 或 MF15398。

没有发现有任何支系存在山西梁姓在全国同姓占比高的情况（附件表一），提示当年的梁国梁姓 MF14287/MF15398/MF14963 族群可能是穿过山西迁徙到了河北，也说明山西不存在任何山西梁姓相较全国梁姓特别高频的支系。总体来说，梁姓与史书记载的一致性进一步支持了黄姓和徐姓的基因学结论。这种在不同姓氏研究中所表现出的一致性，进一步强化了基因数据能够为这些家族的历史起源提供宝贵见解的论点。

与梁姓共享最近祖先支系 MF14287 的现生人包括湖南湘阴的左氏家族，其遗传标记是较末端的一个支系 MF59804 （图二）。属于这个支系的魔方用户有 5 人，除去近亲关系 1 人，有效人数是 4 人，其中 75%都是来自湖南的左姓人士。 晚清中兴四大名臣之一的民族英雄左宗棠老家是湖南湘阴，因此，他较有可能是属于这个支系。

其它嬴姓姓氏

我们也分析了其它两个人口数较多的嬴姓姓氏的基因型分布，赵姓和马姓，但没有发现姓氏高频支系与史书记载聚居地的重合现象。这可能与这些姓氏被记载与伯益的关系较远有关。赵姓和马姓的始祖都是造父，史书记载造父是伯益的 14 世孙。另外，较近代有个赵姓皇帝（宋朝赵匡胤），其后代人数会较多，也可能会造成赵姓族谱的改变或混乱。其它人口较少嬴姓姓氏，暂无法进行有意义的分析，需有待未来有更多魔方客户信息后再进行分析。另外，这些姓氏都与伯益的关系较远或确切性较差，对于验证史书记载的作用也就比较有限。

讨论：

本研究利用 23 魔方基因数据库并结合独立测试和赞助测试，系统分析了黄姓和其它嬴姓相关姓氏人群在不同父系单倍型支系的分布情况，本研究来自《史记》等历史文献的启发，并得出了与《史记》记载基本一致的遗传学结论。

综合文献以及目前黄姓的人口分布，与黄姓起始人群最相关的黄姓人群应该主要分布在湖北地区，本研究分析结果显示，这批人群的祖先可能是 3420-4320 年前的、带有 MF14296/MF15137 基因的某以黄为姓的黄国公族家族(图二)。

如果《史记》等文献关于黄姓源流的记载属实，则下列推论应该成立。1，黄国应该存在，湖北黄姓族群因为来自黄国，理应是最符合黄国记载的黄姓人群，黄姓的形成时间应该是不晚于黄国灭亡时间，即2672 年前。2，伯益及大廉等东夷家族应该存在，这些上古黄夷族部落的人物以及其祖先如黄帝等应该具备超级祖先特征，即其所属父系支系具备后代支系多的特征。3，该家族属于夏朝人物，黄国也是建立于夏朝，故夏朝应该存在。4， 有明确记载的其它嬴姓所属姓氏如徐姓和梁姓等也应该属实，并在基因型上与黄姓所属基因型存在密切关系。

黄国的存在有考古实证，位于河南潢川，但具体起始时间是否在夏朝，仍尚未通过考古考证(8)。本研究发现黄姓起始支系的形成时间在 3420-4320 年前，也就是黄国前期或中期。我们能在湖北发现一个符合史书记载的黄姓支系 MF14296/MF15137，可能不是偶然，而是反映了史书一定的客观性。作为对照，我们未能在一个不相关的省份如江苏做出类似发现。

我们研究虽然发现了湖北黄姓也存在一个 C-MF1017 支系，及其下游支系 C-Z45203，这两个支系也具备湖北黄姓在全国同姓氏中占比较高的特征。但这些支系在其它方面不如 MF14296/MF15137 支系更加符合黄姓始祖应该有的特点，包括较低的黄姓在全国所有姓氏人数的占比、支系的湖北人数在全国的占比较低、和黄姓样本大部分来自距离黄国故城较远的黄石。

因为大量改姓或认祖归宗的赐姓，某古老姓氏的单倍型就不能通过该姓氏在其所属单倍型的全国占比来确定。姓氏和姓氏所属单倍型一般具有地域特异性， 湖北黄姓与单倍型 MF14296 的湖北特异性联系也是代表了姓氏单倍型的地域特征，这种支系和姓氏的区域特异性给姓氏的支系属性研究增加了复杂性，不能简单用全国占比来直接推导。同时也说明，对于某姓氏的基因支系属性，特别是数千年的古老姓氏，只能是区域特异地进行研究才会有意义。如果某姓氏和其所属单倍型源自某个地区省市，则该地区的该姓氏在全国同姓氏同单倍型的占比应该是最高。

黄氏宗亲网列出了很多黄姓家族富集的各类单倍型 （http://www.ihuang.org/a4.htm）。本研究也发现，黄姓也在很多其它单倍型富集，但这些基因型均与文献记载的黄姓源流不一致，从时间和地点来看，只有湖北黄姓单倍群 MF14296/MF15137 最符合史书和考古关于黄国及黄姓的记载。因此，这些不同基因支系的黄姓人群应该不是来自某一个姓黄的最近共同祖先，而有可能是独立选择黄姓作为姓氏。多种复杂因素，如改姓，赐姓、抱养，认祖归宗等等，会抹平源流的印记。黄姓虽然在几乎所有较大支系都有分布，但显然只有其中一个支系才是黄国黄姓的代表。结合姓氏与基因地域特异性的研究，并综合历史文献的考证，可能是比较有效地姓氏遗传学新途径。

本研究从分子遗传学角度独立支持了以上按照文献得出的四个推论。湖北黄姓支系 MF14296 的祖上基因型是 F2137、F8、和 M117，属于三大超级祖先基因型之一（图一 A），也是与黄夷族部落和黄帝记载最为契合的基因型 (6)。MF14296 的先后不同的祖先支系绝大部分都具备超级祖先的特征，比如该支系的最近上游祖先 A16636 具备 6 个下游支系，比同级其它 4 个姊妹支系都多（图一 B），A16636 最近上游支系A16635 具备 5 个下游支系，也比同级其它 3 个姊妹支系都多。总起来看，在 MF14296 到 M117 一共 13 个祖先支系中，只有一个支系 F438 不是同级姊妹支系中最具备超级祖先特征的，其含有 3 个下游支系，次于 F14249（10 个下游支系）、Z25921（4 个下游支系）、和 Z25928（4 个下游支系），但是，F438 后裔人数最多，是 F14249 后裔的 2.91 倍、Z25921 后裔的 2.8 倍、和 Z25928 后裔的 4.2 倍。这提示，MF14296 黄姓支系的各级不同时期的祖先绝大部分具备帝王特征，这显然符合文献里关于黄姓祖先是夏朝重臣伯益，以及伯益可溯源到黄帝和黄夷族部落的记载。同时，本研究通过鉴定黄姓的祖先单倍型为F8/M117，也进一步论证了该单倍型有可能是黄帝的单倍型，因为黄姓祖先被文献记载为是出自上古黄夷部落，部落头领被称号黄帝。

A16635 是 F2137 后裔的最主要支系，今天的 F2137 后裔人群中有 73.2%属于 A16635。古 DNA 研究显示，F2137 在 4000 年前是中原和周边地区遗址的常见单倍型，这些遗址包括甘肃磨沟 4000 年前(9)，陕西石峁 4000 年前(10)，河南渑池仰韶村 4000 年前(11)，内蒙古赤峰大山前 3500 年前(9)。这进一步支持 F2137 是超级祖先的推论，以及文献里关于黄姓及嬴姓祖先是族群头领人物的记载。

本研究显示，有较确切记载的、伯益其它子孙的姓氏，如徐姓和梁姓，也都存在类似湖北黄姓的区域性支系富集特征，徐国徐姓始祖支系可能是在江苏上海较高频的 MF38096/SK1726，梁国梁姓可能是在河北较高频的 MF14287/MF15398/MF14963，他们的基因型和始祖出现时间都基本符合史记的历史记载，他们之间的基因进化关系以及与湖北黄姓的基因进化关系也都符合《史记》等文献记载（图二）。另外，徐姓和梁姓也都被发现有在多个其它支系富集的现象，但这些支系都不是最具备徐姓或梁姓始祖特征的支系。因此，嬴姓所属的 3 个主要姓氏，其中包括两个记载最明确、与伯益关系最紧密的姓氏，即黄姓和徐姓，都在基因层面和地区分布层面支持了《史记》等文献的叙事。

其它嬴姓姓氏的基因型属性还需要未来的深入研究才有可能确定。赵姓和马姓被记载有一个共同祖先造父，一个距离伯益较远的伯益后人。本研究未能鉴定出他们的始祖基因型，这除了与伯益关系较远或较不确定有关，也可能与较近代有个赵姓皇帝（宋朝赵匡胤）有关。皇帝影响下的改姓、赐姓等等因素可能会冲淡姓氏的远古源流特征。在封建社会，避讳是一种普遍现象。人们为了避免与皇帝或尊长的名字相同或相似，可能会选择改姓或改名。另外，在宋朝灭亡后，一些赵姓为了避免元朝的迫害，选择了改姓。

《史记》所提及的嬴姓 14 氏的其他姓氏因为人数少，暂未能被包括在本研究中，另外，他们与伯益的关系也在古籍中较不明确，并且其中很多与现代姓氏的对应关系也较不明确，因此，这些姓氏的研究对于史书记载的独立验证意义比较有限。

如果最初黄姓来自黄国国民以国为姓，并且黄国国君也是黄姓并带有某单倍型 X，则可以预期，黄国国民或黄姓绝大部分应该不带有单倍型 X，因为即使国君家族的后裔人数会大于一个普通百姓家族，但黄国被封给国君时，理应已经有大量人口，这些人口的父系基因分布，应该反映了当时的中原地区不同父系人口的分布。另外，也可能会有大量其它外地人口加入黄国，这也会增加黄国居民的基因多样性。这种以国名为姓的传统，必然会导致大量的、具备不同基因型的人来共享同一个姓氏。

黄姓和徐姓的始祖支系分属两个不同支系，A16636 和 F15823，都是 A16635 的最近下游支系，一种简单直观逻辑会把伯益归为 A16635，而他的两个儿子则分属 A16636（大廉）和 F15823（若木）。但这种关系虽然不是不可能，但也可能存在有其它关系。因为 Y 染色体突变率是每 84.5 年一个突变（23 魔方标准），伯益与其儿子的 Y 染色体有可能一样但也有可能不同。若一样，则他们三人有可能都是 A16635，A16636 的突变可能发生在大廉的某个后裔的身上，而 F15823 可能发生在若木的某个后裔身上。

A16635 的共祖时间（5340-5360）与其下游 A16636 和 F15823 各自的共祖时间存在较大差别（分别是 4330-5340 和 5230-5340），A16635 与 A16636 有可能存在近 1000 年的差别，这提示 A16635 与这个大廉支系的关系不像是父子关系，这也支持 A16636 的突变应该不是发生在伯益儿子大廉这一代，而是发生在更远的大廉的后裔个体中。

但需要注意的是，这相隔 1000 年的差别有可能不是真实差别，实际差别可能会小较多，因为重复突变其实很常见，新近的古 DNA 研究和 Y 染色体完整测序研究都论证了一个新理念和现象，很多被认为是祖先支系发生的突变，其实都是来自不同的下游后代支系的各自的趋同独立突变(12, 13)。因此，年代越久的祖先支系就会包含越多的来自后代的趋同突变，这些突变自然会增加祖先支系与最近后代支系的年代

差别。由于仅仅平均 84.5 年就会产生一个突变，一个祖先支系与其最近下游支系的时间差别最多应该是在 84 年到几百年内。因此，1000 年的差别基本是不太可能的，这只能说明这种根据突变数目来推算时间的方法有局限性，没有考虑重复突变的因素。

这也可以解释为何同级两个支系，A16636 支系有多个突变（9 个突变,根据国际基因族谱学会 ISOGG 网站数据）或距离祖先 A16635 年代更久，而 F15823 支系却只有很少的突变（仅一个突变）或距离 A16635 非常近。这也独立支持了父系染色体重复突变非常普遍的这一最新发现。总之，目前魔方关于 A16635 的共祖年代是 5300 年的推算，仅是提供了一种近似的参考，与实际相差或小于实际数百年乃至千年都是可能的。因此，目前推算的 A16635 等主要支系的共祖年代，与文献记载的相关人物出现的年代（约 4000-5000 年前后）并没有显著的矛盾。另外，F8 被 23 魔方推算共祖时间为 7290-12910 年，也同样仅是一个误差度较高的估算，并与其它学者推算的时间（约 5400 年）有较大出入 (6)。

《史记·秦本纪》提到： “大费（即伯益）生子二人：一曰大廉，实鸟俗氏；二曰若木，实费氏。”又提到：“大廉玄孙曰孟戏、中衍，鸟身人言。帝太戊闻而卜之使御，吉，遂致使御而妻之。自太戊以下，中衍之後，遂世有功，以佐殷国，故嬴姓多显，遂为诸侯。” 这说明大廉的后代中，有名为孟戏和中衍的玄孙（大廉的第五代子孙）， 长着鸟的身子但会人说人话（这虽然可能带有神话色彩，但也反映了古人对大廉后代的某种神秘或特殊的想象），太戊听说后占卜了一下，结果吉，于是就派孟戏担任御者，并把女儿嫁给他。从太戊以下（约公元前 1401 年），由于中衍的后代，多个世代都有佐助殷国的功绩，所以嬴姓的人大多显贵，于是成为诸侯。这直接说明了大廉的后代在商朝时期有着较高的地位和影响力，同时也说明，大廉确实传有男性子孙，符合黄国国君一族世袭多代超过千年的记载，也支持黄国在商朝时就已经存在。

“嬴性多显”能被历史学者注意到并写入史书，可能与这个特征的不同寻常有关。本研究提示，这可能与遗传因素有一定关系，因为不仅古代商朝时是这样（如史书记载），上古时期也可能是同样，因为嬴姓上游基因支系基本都具备超级祖先特征，特别是包含了最有可能是黄帝所属的支系 M117-F8-F2137。统一中国的核心人物秦始皇名叫嬴政，他可能属于嬴姓伯益的后裔，但具体传承关系在历史上并没有明确的文献记载,需要未来深入研究来确定。另外，较近代的嬴姓后裔中也可能出现过伟人。通过鉴定嬴姓始祖基因型，就让发现属于这个基因型的近代名人伟人成为可能，湖南湘阴左氏家族就带有这个基因型（图二），近代民族英雄左宗棠有可能是源自这个家族，但这仍然需要进一步研究来确认。M117-F8 支系在数千年左右的时间内发展到今天约 17%的汉族男性，从后裔人数上在新时期时代的超级祖先中排名第一(6)，这一现象也支持这一支系的强大生存竞争能力，但是否能在父系染色体上找到直接与生存能力有关的基因，还需要未来的深入研究来决定。现有的关于复杂性状的遗传研究，如 GWAS（全基因组关联研究）类型的研究，通常会忽略来自 Y 染色体的遗传变异。

本研究的遗传学数据表明，大廉本人以及他较近的子孙可能属于 A16635 或 A16636,他们可能并未正式采用“黄”作为姓氏。相反，黄姓的正式使用可能是源自某个携带 MF14296 遗传标记的大廉后裔个体（如图二所示），这一变化大致发生在大廉创建黄国后的数百年间。《史记》所记载的大廉第五代孙的名字是孟戏和中衍，并非姓黄。 同样地，徐姓的正式启用也并非直接源自若木本人或其较近子孙，而是源自某个携带 MF38096 或 SK1726 遗传标记的若木后裔个体（如图二所示），这一变化大致发生在若木创建徐国后的数百年间。此前，有一种观点认为，姓氏的正式、广泛使用始于周朝（公元前 1046 年至公元前 256 年），这一时期的宗法分封制度对姓氏的普及起到了关键作用。尽管在周朝之前，姓氏的萌芽形态已存在，但其使用可能并不普遍 (8)。本研究的结论与上述观点并无显著冲突。实际上，它表明黄姓和徐姓的正式使用可能略早于周朝，但仍然是在宗法制度逐渐成形并影响社会结构的大背景下发生的。这一发现进一步丰富了我们对姓氏起源和演变历程的理解。

关于黄姓始祖的问题，文献上存在不同的说法。其中一种主要观点认为颛顼（黄帝之孙）的后裔姬姓陆终是黄姓始祖，这一观点最早出可见于南朝（公元 420-589 年）著名学者王俭所著的《姓谱》，这与大廉为黄姓始祖的叙事看似存在冲突，但实际上可能反映了不同历史时期和文化背景下的认知差异。为了调和这一矛盾，学者们提出了多种猜想。其中一种猜想是，陆终可能代表的是母系社会时期的姬姓始祖，而伯益家族则代表的是进入父系社会后的嬴姓始祖(14)。本研究的结论较符合这一猜想，如果陆终是黄姓父系始祖，则陆终黄姓就不会与徐姓或梁姓有亲缘关系，因为文献里没有出现过相关记载。但本研究所发现的黄姓与徐姓及梁姓的遗传关系符合《史记》中关于伯益的族谱叙事记载（图二），这就一定程度上化解了不同文献中关于黄姓记载之间的冲突，并倾向于支持大廉是黄姓父系始祖的结论,这也进一步说明了司马迁在其伟大作品《史记》中所展现出的深厚学术造诣。

《史记》等文献中有大量关于夏朝的记载，但鉴于西方学者的质疑与标准，以及考古证据的暂时缺乏，学界关于夏朝是否存在还未达成一致意见。新的独立研究角度有可能为夏朝的验证提供独特证据视角。黄国和徐国都被记载是建立于夏朝初期，是夏朝重臣伯益的两个儿子的封地，并导致了黄姓和徐姓的起源。本研究从姓氏遗传学角度溯源了这两个姓氏的基因型起源，得到了基本符合黄国和徐国历史记载的遗传学结论，这也就为这两个国家和夏朝的存在提供了强有力的遗传学佐证。

尽管本研究的总体样本量颇为可观，但当细化到个别姓氏时，其样本量则显得相对有限。鉴于此，本研究的结论有望通过未来更大规模、更具针对性的系统研究得以进一步巩固。然而，值得注意的是，本研究所得结论已具备较高可信度，原因在于，三项针对不同姓氏的独立研究均得出了与文献记载相一致且相互印证的结论。倘若较小的样本量确实导致了结论的偏差，那么这种跨姓氏的一致性便难以维系。此外，本研究采用的样本数据绝大部分源自 23 魔方，该公司的样本收集过程并未针对特定姓氏进行特殊处理，从而确保了研究结论较少受到片面性的影响。

总之，本研究鉴定了与文献记载的黄姓源流基本符合的一支湖北黄姓基因支系，也同时鉴定了黄姓相关姓氏苏沪徐姓和河北梁姓的基因支系。这些支系的现实存在显示了文献记载的客观性，佐证了夏朝初年建立的黄国和徐国的存在。通过姓氏遗传学来独立验证文献中关于姓氏起源与流传，以及相关朝代、国家、和历史人物的记载的真实性，可能会在未来产生很多富有意义的分子历史研究新成果。

致谢

我们感谢刘小舟和夏志毅的技术支持。本研究得到了中国国家自然科学基金 81171880 的支持。

利益冲突：
作者承诺，本研究没有与任何方面有利益冲突。

作者贡献：

黄石参与了研究设计，数据分析，指导团队技术工作，和论文撰写。

表格：

表一，23 魔方数据库有关姓氏在湖北和广东两省的部分单倍型支系的分布，分析截止日 2024 年 11 月 18 号

单倍型	形成时间	姓氏	汉族人数全国		全国	单个姓氏人数		某省某姓全国同姓占比	
	年前		单个姓氏	全部姓氏	姓氏比例	湖北	广东	湖北	广东
A0-T	149960	黄	4957	207333	0.024	234	1159	0.047	0.234
		陈	11529		0.056	446	1776	0.039	0.154
M117	13870	黄	993	39381	0.025	40	258	0.040	0.260
		陈	2361		0.060	73	439	0.031	0.186
F8	7250	黄	934	37064	0.025	36	261	0.039	0.279
		陈	2267		0.061	70	427	0.031	0.188
ACT2839	7250	黄	934	35923	0.026	35	250	0.037	0.268
F438	6800	黄	184	12690	0.014	13	29	0.071	0.158
F14366	6790	黄	166	11857	0.014	13	27	0.078	0.163
Y17728	6610	黄	159	11357	0.014	13	24	0.082	0.151
ACT7481	6590	黄	60	5052	0.012	7	14	0.117	0.233
F1754	6530	黄	57	4927	0.012	7	14	0.123	0.246
F26655	6510	黄	2	295	0.007	0	0	0.000	0.000
F2643	5630	黄	52	4509	0.012	7	14	0.135	0.269
F2137	5350	黄	52	4509	0.012	7	14	0.135	0.269
A16635	5330	黄	44	3110	0.014	7	11	0.159	0.250
		陈	113		0.036	3	24	0.027	0.212
		王	233		0.075	7	10	0.030	0.043
		张	251		0.081	7	19	0.028	0.076
		李	198		0.064	11	16	0.056	0.081
		刘	319		0.103	9	64	0.028	0.201
		周	81		0.026	1	18	0.012	0.222
		杨	74		0.024	2	4	0.027	0.054
		徐	38		0.012	0	5	0.000	0.132
		吴	49		0.016	3	3	0.061	0.061
		赵	76		0.024	2	1	0.026	0.013
		朱	43		0.014	2	6	0.047	0.140
A16636	4320	黄	6	735	0.008	3	0	0.500	0.000
		陈	17		0.023	0	1	0.000	0.059
		李	69		0.094	3	1	0.043	0.014
MF14296	3580	黄	6	56	0.107	3	0	0.500	0.000
MF607191	2780	黄	4	41	0.098	3	0	0.750	0.000
		陈	1		0.024	0	0	0.000	0.000
F15823	5190	黄	38	2479	0.015	3	11	0.079	0.289
		陈	96		0.039	3	23	0.031	0.240
		李	129		0.052	8	15	0.062	0.116
MF15397	4980	黄	24	1869	0.013	3	9	0.125	0.375
		陈	73		0.039	2	21	0.027	0.288
		李	86		0.046	5	11	0.058	0.128

表二，黄姓、徐姓、和梁姓在各个单倍型支系不同地区的占比

全国样本数少于 5 的支系没有列出。在列出的支系中，如果某姓氏的全国样本数少于 5 人，则就没有显示样本数的数据。

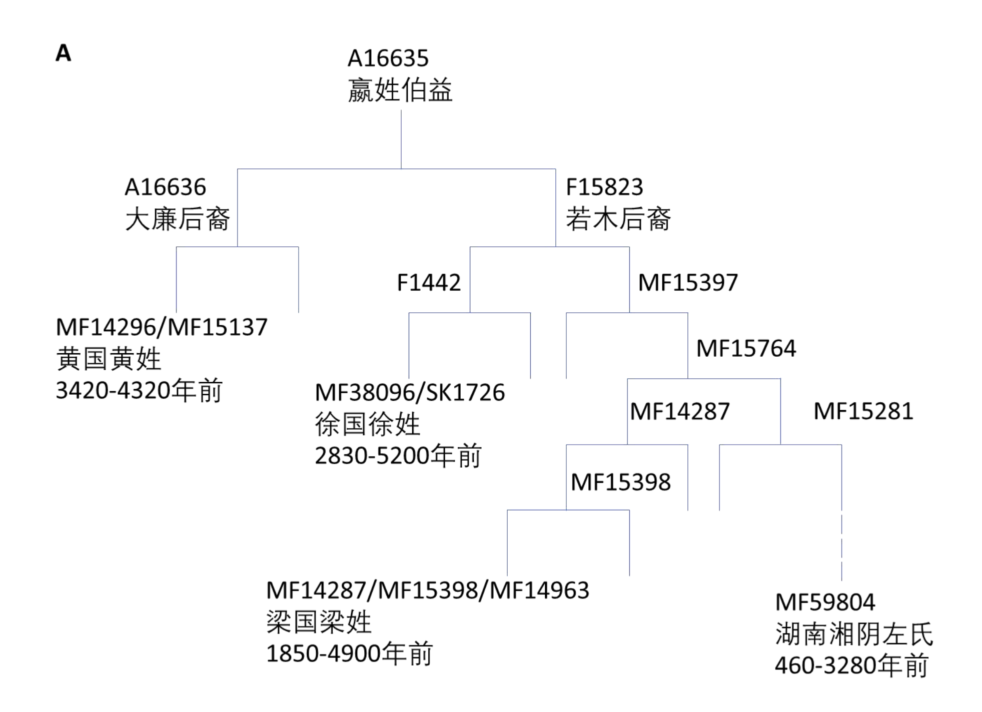
单倍型	形成年	全国黄人数	湖北黄占比	全国徐人数	苏沪徐占比	全国梁人数	河北梁占比	全国占比					
								黄姓	徐姓	梁姓	湖北	苏沪	河北
人口普查			0.042		0.13			0.022	0.014	0.007	0.043	0.076	0.054
O-MF14296	3580	6	0.50					0.11			0.20		
C-MF1017	3760	19	0.53					0.05			0.1		
C-Z45203	2980	14	0.57					0.09			0.084		
O-MF38096	3370			5	1.00				0.1			0.25	
O-F1442	5200	13	0.08	10	0.70				0.018			0.13	
O-B432	4090			8	1.00				0.039			0.30	
O-MF1198	4770	9	0.00	14	0.57				0.018			0.38	
O-MF941	4860			5	0.60				0.025			0.15	
O-MF6935	3440	5	0.20	7	0.57				0.012			0.12	
O-MF6934	5330	14	0.07	20	0.60				0.021			0.17	
O-MF1075	4450	18	0.11	25	0.72				0.02			0.38	
O-MF6080	4450	26	0.00	6	0.67				0.011			0.18	
O-PH4822	3460			8	0.50				0.036			0.19	
O-CTS5664	4570			5	0.80				0.017			0.087	
O-Z17557	3500	11	0.00	5	0.60				0.015			0.34	
O-CTS679	3570	12	0.00	6	0.50				0.012			0.079	
O-MF18577	3960	5	0.00	19	0.58				0.032			0.42	
O-MF1107	4130	17	0.12	6	0.67				0.028			0.34	
O-MF182502	4740	5	0.00	23	0.61				0.033			0.40	
O-MF9896	5020			6	1.00				0.057			0.12	
O-Y89945	5190	6	0.00	5	0.60				0.015			0.43	
O-MF14135	5280	9	0.00	41	0.54				0.04			0.32	
N-FT317554	4480			5	0.60				0.028			0.19	
N-F839	5050	5	0.20	8	0.63				0.023			0.23	
C-F10085	4230	6	0.00	5	0.60				0.0092			0.047	
C-SK1038	4800	37	0.38	8	0.50				0.007			0.030	
C-MF10317	5320			8	0.75				0.04			0.22	
O-MF14287	4090					5	0.80			0.1			0.19
O-MF15398	3210					5	0.80			0.11			0.20
O-MF14963	1850					5	0.80			0.11			0.205
D-F2625	5450	5	0.20			5	0.60			0.0082			0.056
O-CTS723	3190	5	0.20	5	0.40	5	0.60			0.0073			0.075

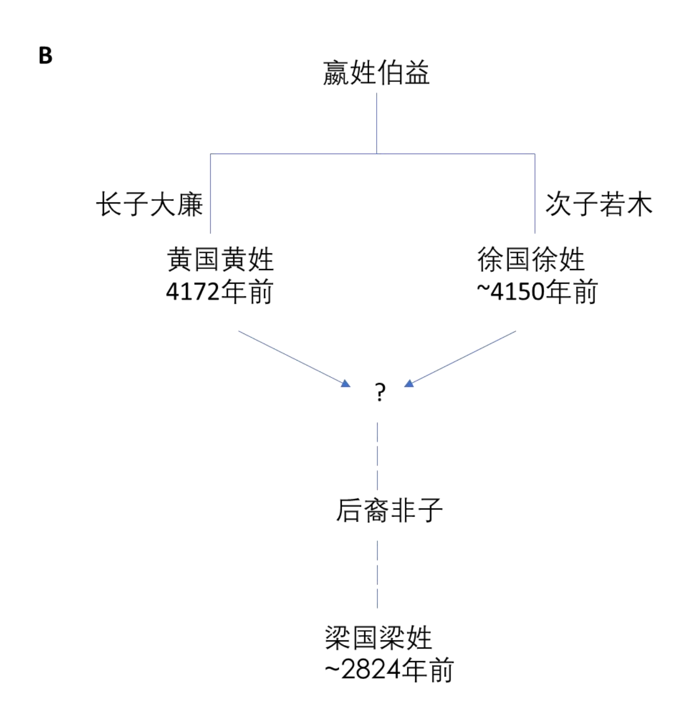
图文字说明

图一，湖北黄陂罗汉寺黄姓家族某黄姓个体 MF287231 的父系支系进化树

进化树结构来自 23 魔方。相关祖先支系的遗传标记如图所示在支系线的最近上方，每个支系的后代支系数目也如图所示。姊妹支系按照后代支系数目多少从多到少自上而下排列。A，从 M117 到 A16636 的支系树。B， 从 A16635 到 MF287231 的支系树。

图二，嬴姓相关姓氏的谱系树。 A）嬴姓相关姓氏的基因型族谱树。B）《史记》等文献记载的嬴姓相关姓氏的谱系树。





References:

1. Yuan YD. Science and culture of surnames. Chinese Nat Geog. 2007;2:38-9.
2. Liu Y, Chen L, Yuan Y, Chen J. A study of surnames in China through isonymy. Am J Phys Anthropol. 2012;148(3):341-50.
3. Jobling MA. In the name of the father: surnames and genetics. Trends Genet. 2001;17(6):353-7.
4. Yuan YD, Qiu JR. Dictionary of Chinese Surnames. Jiangxi, China: Jiangxi People's Press; 2010.
5. 番茄汁. 黄冈、黄石、黄陂……为什么湖北有那么多姓"黄"的地名？. 澎湃新闻. 2021:https://m.thepaper.cn/baijiahao_9809756.
6. Yan S, Wang CC, Zheng HX, Wang W, Qin ZD, Wei LH, et al. Y chromosomes of 40% Chinese descend from three Neolithic super-grandfathers. PLoS ONE. 2014;9(8):e105691.
7. 黄锁奎. 黄姓寻根探秘：认祖归宗以国为姓. 香港: 香港黄氏寻根编委会; 2016.
8. 黄展岳. 黄姓源头——青铜器里的潢川黄国. 大众考古. 2015;9:29-32.
9. Zhang Y, Lei X, Chen H, Zhou H, Huang S. Ancient DNAs and the Neolithic Chinese super-grandfather Y haplotypes. bioRxiv. 2018;doi: https://doi.org/10.1101/487918.
10. Ning C, Li T, Wang K, Zhang F, Li T, Wu X, et al. Ancient genomes from northern China suggest links between subsistence changes and human migration. Nature communications. 2020;11(1):2700.
11. Li S, Wang R, Ma H, Tu Z, Qiu L, Chen H, et al. Ancient genomic time transect unravels the population dynamics of Neolithic middle Yellow River farmers. Sci Bull (Beijing). 2024;69(21):3365-70.
12. Xia Z, Chen H, Zhang Y, Huang S. Ancient uniparental DNAs in distinguishing the competing theories of molecular evolution and modern human origins. Research Square. 2021:DOI:10.21203/rs.3.rs-1087730/v1.
13. Huang S. Examining models of modern human origins through the analysis of 43 fully sequenced human Y chromosomes. BioRxiv. 2023(doi: https://doi.org/10.1101/2023.11.09.566475).
14. 黄文节, 杨文祥. 《我的家》上篇 寻根——黄氏宗族溯源. 2020:https://blog.sciencenet.cn/blog-496942-1253150.html.

Supplementary Table S1. 黄姓、徐姓、和梁姓在各个单倍型支系不同地区的占比

全国样本数少于 5 的支系没有列出。在列出的支系中，如果某姓氏的全国样本数少于 5 人，则就没有显示样本数的数据。

单倍型	形成年	湖北黄人数	全国黄人数	比例	苏沪徐人数	全国徐人数	比例	河北梁人数	山西梁人数	全国梁人数	河北梁比例	山西梁比例	江苏黄	比例	全国占比					
															黄姓	徐姓	梁姓	湖北	苏沪	河北
人口普查				0.042			0.13							0.042	0.022	0.014	0.007	0.043	0.076	0.054
M117																				
F440	3050	5	74	0.07	16	69	0.23	1	0	17	0.06	0.00	3	0.04						
MF38096	3370				5	5	1.00									0.1			0.25	
F14952	3470				0	6	0.00													
MF14296	3580	3	6	0.50									1	0.17	0.11			0.20		
SK1730	3640	0	95	0.00	0	21	0.00	1	0	80	0.01	0.00	3	0.03						
F1128	3670				2	6	0.00													
M1528	3790	2	154	0.01	6	34	0.18	1	0	8	0.13	0.00	1	0.01						
PH3164	3820				0	14	0.00													
CTS5492	3930	4	151	0.03	2	52	0.04	0	0	11	0.00	0.00	13	0.09						
MF14287	4090							4	0	5	0.80	0.00					0.1			0.19
MF50069	4230				2	8	0.25													
A16636	4320	3	6	0.50	1	9	0.11						1	0.17	0.007					
F1707	4370	0	49	0.00	2	11	0.18	0	0	11	0.00	0.00	0	0.00						
F6353	4460	0	8	0.00									0	0.00						
CTS466	4530	1	9	0.11									3	0.33						
MF2604	4700	1	24	0.04	4	13	0.31						7	0.29						
MF16648	4860	0	8	0.00	1	5	0.20						0	0.00						
YP4864	4950	1	14	0.07	4	26	0.15						0	0.00						
F15823	5190	3	38	0.08	13	28	0.46	9	1	21	0.43	0.05	1	0.03						
F1442	5200	1	13	0.08	7	10	0.70						0	0.00		0.018			0.13	
MF10031	5230	0	8	0.00	1	10	0.10	1	0	10	0.10	0.00	0	0.00						
F310	5240	1	41	0.02	11	31	0.35	0	0	9	0.00	0.00	4	0.10						
MF61476	5250	0	11	0.00	0	9	0.00	1	2	8	0.13	0.25	0	0.00						
A16635	5330	6	44	0.14	14	37	0.38	9	0	24	0.38	0.00	3	0.07						
F2137	5350	6	52	0.12	20	49	0.41	12	2	30	0.40	0.07	4	0.08						
CTS5308	5370	0	11	0.00	1	11	0.09	1	0	11	0.09	0.00	0	0.00						
MF15246	5380				5	15	0.33													
F316	5410	5	72	0.07	16	71	0.23	1	0	21	0.05	0.00	3	0.04						
F46																				
F4249	3000	7	35	0.20	16	52	0.31	0	2	6	0.00	0.33	0.00	0.00						
Z26297	3540	4	10	0.40									1.00	0.10						
MF300768	3550				0	6	0.00													
MF357	3620	0	25	0.00	7	15	0.47	0		4	0.00		0.00	0.00		0.032				
B432	4090				8	8	1.00	0		2	0.00					0.039			0.30	
MF25870	4130				0	5	0.00													
Z26276	4360	1	21	0.05	5	14	0.36						0.00	0.00						
Z26258	4390	1	24		6	24	0.25	0	0	9	0.00	0.00	1.00	0.04						
MF91	4460				0	6	0.00	0	0	5	0.00	0.00								
F152	4640	0	25	0.00	5	22	0.23	1	0	6	0.17	0.00	1.00	0.04						
F4117	4750				1	7	0.14													
MF1198	4770	0	9	0.00	8	14	0.57	0		2	0.00		2.00	0.22		0.018			0.38	
Z26246	4780	2	25	0.08	6	25	0.24	1	0	9	0.11	0.00	1.00	0.04						
F14839	4850	0	17	0.00	12	40	0.30	2	0	12	0.17	0.00	2.00	0.12						
MF941	4860				3	5	0.60									0.025			0.15	
Z26248	5020	2	27	0.07				0	0	9	0.00	0.00	1.00	0.04						
Z26108	5230	1	32	0.03	12	33	0.36	0	0	8	0.00	0.00	5.00	0.16						
F11																				
F793	2730	5	70	0.07	3	24	0.13	1	0	19	0.05	0.00	1.00	0.01						
MF6935	3440	1	5	0.20	4	7	0.57						0.00	0.00		0.012			0.12	
MF1107	4130	2	17	0.12									4.00	0.24						
A22321	4300	0	8	0.00	2	7	0.29						1.00	0.13						
MF1309	4340	0	10	0.00				0	0	8	0.00	0.00	0.00	0.00						
MF6739	4420	0	15	0.00	7	20	0.35	1	0	6	0.17	0.00	2.00	0.13						
F2941	4440	0	23	0.00	6	19	0.32	0	1	5	0.00	0.20	5.00	0.22						
F12	4480	1	67	0.01	9	40	0.23	0	1	7	0.00	0.14	12.00	0.18						
F930	4550	1	41	0.02	4	25	0.16	0	0	8	0.00	0.00	1.00	0.02						
Z43876	4580	5	72	0.07	3	24	0.13	0	0	19	0.00	0.00	1.00	0.01						
MF6915	4600	0	8	0.00									1.00	0.13						
CTS7789	4610	2	37	0.05	1	12	0.08	0	0	6	0.00	0.00	1.00	0.03						

Y29838	5050	3	23	0.13	2	14	0.14						3.00	0.13						
FGC54486	5050	2	19	0.11	8	22	0.36	2	1	7	0.29	0.14	1.00	0.05						
Y141175	5220	5	76	0.07	3	24	0.13	1	0	19	0.05	0.00	1.00	0.01						
MF161232	5310				2	10	0.20													
F386	5310	1	8	0.13	2	7	0.29						2.00	0.25						
MF6934	5330	1	14	0.07	12	20	0.60						3.00	0.21		0.021			0.17	
M119																				
CTS11040	3210	3	13	0.23	0	6	0.00						0	0.00						
MF6285	3280				2	10	0.20													
BY47757	3650				2	5	0.40													
CTS9421	3710	0	22	0.00									0	0.00		0.0079				
F970	4060	0	11	0.00	2	31	0.06	0	0	5	0.00	0.00	1	0.09						
CTS716	4070	2	28	0.07	1	11	0.09	0	0	9	0.00	0.00	1	0.04						
CTS8501	4220	8	306	0.03	101	304	0.33	0	1	100	0.00	0.01	39	0.13						
MF1075	4450	2	18	0.11	18	25	0.72						5	0.28		0.02			0.38	
MF6080	4450	0	26	0.00	4	6	0.67						0	0.00		0.011			0.18	
M101	4590	0	14	0.00	1	9	0.11						0	0.00						
MF6180	4640	0	17	0.00	7	29	0.24	0	0	5	0.00	0.00	4	0.24						
FGC66058	4680	3	18	0.17	11	131	0.08	0		2	0.00		2	0.11						
Z23392	4760	2	69	0.03	7	20	0.35	1	0	19	0.05	0.00	3	0.04						
ACT263	4850	0	6	0.00	1	14	0.07						3	0.50						
MF6122	5000	3	14	0.21	0	7	0.00	0	0	5	0.00	0.00	0	0.00						
Z23271	5040	0	20	0.00				0	0	10	0.00	0.00	0	0.00						
F1036	5350	0	11	0.00	2	34	0.06	0	0	5	0.00	0.00	1	0.09						
M268																				
CTS723	3190	1	5	0.20	2	5	0.40	3	0	5	0.60	0.00	0	0.000			0.0073			0.075
PH4822	3460				4	8	0.50									0.036			0.19	
Z25482	3500	2	121	0.02	0	14	0.00	0	1	7	0.00	0.14	1	0.008						
CTS1127	3740	0	7	0.00	2	16	0.13						0	0.000						
Z33440	3900	0	10	0.00	1	19	0.05						0	0.000						
MF16487	3910	0	16	0.00	2	22	0.09						0	0.000						
F3323	3940	1	86	0.01	22	117	0.19	3	1	19	0.16	0.05	4	0.047						
Y10774	4360	0	12	0.00	2	15	0.13						2	0.167						
L682	4450	1	8	0.13	2	5	0.40	1	0	5	0.20	0.00	1	0.125			0.0072			
CTS6566	4550				0	24	0.00													
CTS5664	4570				4	5	0.80									0.017			0.087	
Z24083	4710	0	84	0.00	2	13	0.15	0	0	51	0.00	0.00	1	0.012						
MF61620	4810	0	8	0.00									0	0.000						
CTS7399	4980	10	46	0.22	2	19	0.11	3	0	44	0.07	0.00	0	0.000						
CTS651	5100	3	47	0.06	1	8	0.13	0	0	9	0.00	0.00	0	0.000						
F18475	5190	0	25	0.00	1	10	0.10						3	0.120						
F840	5290	0	18	0.00	4	19	0.21						3	0.167						
F18990	5400	0	5	0.00	1	6	0.17						1	0.200						
Z24050	5400	3	44	0.07	2	15	0.13	0	0	10	0.00	0.00	0	0.000						
Other O																				
F15433	3320	2	5	0.40	1	5	0.20						0	0.000						
F1262	3490	1	8	0.13	2	9	0.22						0	0.000						
Z17557	3500	0	11	0.00	3	5	0.60						4	0.364	0.028	0.015			0.34	
Z25482	3510	2	121	0.02	0	14	0.00	0	1	7	0.00	0.14	1	0.008						
MF35829	3550	0	7	0.00									0	0.000						
CTS679	3570	0	12	0.00	3	6	0.50						1	0.083		0.012			0.079	
N5	3750	1	13	0.08	0	2	0.00	0	0	10	0.00	0.00	0	0.000						
MF193618	3760	0	20	0.00	0	2	0.00	0	0	5	0.00	0.00	0	0.000						
MF17537	3860	0	20	0.00	0	2	0.00						0	0.000						
SK1783	3910	3	13	0.23	2	12	0.17	1	0	6	0.17	0.00	1	0.077						
MF18577	3960	0	5	0.00	11	19	0.58						1	0.200		0.032			0.42	
F1694	4020	0	18	0.00	12	33	0.36	0	1	5	0.00	0.20	1	0.056						
MF1107	4130	2	17	0.12	4	6	0.67						5	0.294		0.028			0.34	
F728	4190	0	9	0.00	2	14	0.14						0	0.000						
MF57258	4200	0	10	0.00	2	3	0.67						4	0.400	0.024					
Y140772	4220	0	8	0.00	3	22	0.14	0	0	6	0.00	0.00	1	0.125						
F1094	4300	0	9	0.00	0	6	0.00						0	0.000						
CTS6279	4380	0	14	0.00	9	31	0.29	4	0	24	0.17	0.00	1	0.071						

MF7269	4440	0	5	0.00	2	9	0.22						1	0.200						
FGC61169	4470				1	7	0.14													
Y173835	4610	0	14	0.00	4	10	0.40						2	0.143						
MF1240	4640	1	8	0.13	1	3	0.33						0	0.000						
F724	4710				0	6	0.00													
MF182502	4740	0	5	0.00	14	23	0.61						1	0.200		0.033			0.40	
MF1346	4780				0	6	0.00													
MF6592	4790	0	11	0.00	4	12	0.33						2	0.182						
F1022	4860				1	7	0.14													
BY62289	4880	2	11	0.18	4	9	0.44						3	0.273						
F4068	4890	0	11	0.00	8	17	0.47						5	0.455	0.026	0.048				
MF36	5010	0	6	0.00									1	0.167						
Y29792	5010	0	29	0.00	12	41	0.29	0	1	10	0.00	0.10	2	0.069						
MF9896	5020				6	6	1.00									0.057			0.12	
Y89945	5190	0	6	0.00	3	5	0.60						1	0.167		0.015			0.43	
MF14135	5280	0	9	0.00	22	41	0.54						1	0.111		0.04			0.32	
MF15178	5290				2	5	0.40													
F1134	5370	1	8	0.13	2	9	0.22						0	0.000						
MF87763	5480	1	8	0.13									1	0.125						
N																				
MF55354	3700	0	8	0.00									0	0.00						
V1267	3880				2	11	0.18													
F1474	4160				0	5	0.00													
F1437	4280	3	41	0.07	0	6	0.00	2	0	48	0.04	0.00	0	0.00						
FT317554	4480				3	5	0.60									0.028			0.19	
F1154	4530	3	29	0.10	6	48	0.13	0	0	7	0.00	0.00	1	0.03						
CTS1390	4570	0	5	0.00									1	0.20						
MF15765	4690	0	68	0.00	4	14	0.29	0	0	11	0.00	0.00	33	0.49						
Y60861	4750	1	12	0.08									0	0.00						
F1228	4870	0	5	0.00	7	23	0.30						1	0.20						
Y24191	4910	0	5	0.00	0	5	0.00						0	0.00						
Y24191	4910	0	5	0.00	0	5	0.00						0	0.00						
F1525	4970	3	43	0.07	0	6	0.00	2	0	48	0.04	0.00	0	0.00						
F839	5050	1	5	0.20	5	8	0.63						0	0.00		0.023			0.23	
Y67311	5520	0	9	0.00									0	0.00						
C																				
Z45203	2980	8	14	0.57											0.09			0.084		
MF2001	3240				2	6	0.33	0		2	0.00									
MF1648	3270				2	8	0.25													
F10036	3640	1	22	0.05	15	35	0.43	0	0	10	0.00	0.00	1	0.05						
MF1017	3760	10	19	0.53									0	0.00	0.05			0.1		
MF2218	4070	0	31	0.00				0	0	30	0.00	0.00	0	0.00						
MF1718	4120				2	8	0.25													
F3850	4140	0	23	0.00	4	17	0.24	3	1	12	0.25	0.08	3	0.13						
M407	4160	0	23	0.00	4	22	0.18	4	1	16	0.25	0.06	1	0.04						
F10085	4230	0	6	0.00	3	5	0.60						0	0.00		0.0092			0.047	
AM00848	4310	0	13	0.00				0	0	7	0.00	0.00	0	0.00						
F3826	4440	0	5	0.00									0	0.00						
Y88318	4500	0	5	0.00	8	17	0.47						3	0.60						
MF1553	4570	0	12	0.00	3	23	0.13	0	0	6	0.00	0.00	2	0.17						
FGC39603	4600	0	6	0.00									1	0.17						
MF15754	4670	0	5	0.00									2	0.40						
F3830	4760	0	7	0.00	3	10	0.30	0	0	5	0.00	0.00	0	0.00						
SK1038	4800	14	37	0.38	4	8	0.50						0	0.00		0.007			0.030	
MF1777	4890	0	5	0.00									0	0.00						
Z45207	4900							0	1	7	0.00	0.14								
FGC45548	5040	2	15	0.13	9	23	0.39	0	0	7	0.00	0.00	0	0.00						
MF1037	5190				5	13	0.38													
MF1038	5200	0	6	0.00	7	19	0.37						0	0.00						
MF10317	5320				6	8	0.75									0.04			0.22	
A14911	5420	0	6	0.00	4	11	0.36						0	0.00						

D																				
MF2341	2660							0	0	5	0.00	0.00	0							
Z41068	3520	0	13	0.00				0	0	16	0.00	0.00	0	0.00						
F17409	3680	1	6	0.17									1	0.17						
Y14736	3740	0	6	0.00									1	0.17						
MF133667	4070	1	6	0.17	1	5	0.20						1	0.17						
F17412	4980	1	11	0.09	2	8	0.25	1	0	7	0.14	0.00	1	0.09						
F2625	5450	1	5	0.20				3	0	5	0.60	0.00	0	0.00			0.0082			0.056
E																				
L242.1	4780				1	33	0.03													
S9747	5450				2	11	0.18													
Q																				
YP754	3610	0	5	0.00									0	0.00						
Y144860	3870							1	0	5	0.20									
MF1647	3900							2	1	9	0.22									
Y560	4250							1	3	16	0.06									
F1626	4410	4	74	0.05	22	76	0.29	4	0	34	0.12		6	0.08						
R																				
F1345	3480				1	6	0.17													
F3105	4500	0	12	0.00	3	17	0.18	2	2	9	0.22		0	0.00						
Others																				
A		0	0		0	0		0	0	0			0							
B		0	0		0	0		0	0	0			0							
G		0	4	0	0	1	0	0	0	1	0.00	0	0	0						
H		0	1	0	0	0		0	0	0			0	0						
I		0	0		0	0		0	0	0			0							
J		0	15	0	0	7	0	0	2	8	0.00	0.25	0	0						
L		0	0		0	0		0	0	1	0.00	0	0							
M		0	0		0	0		0	0	0			0							
S		0	0		0	0		0	0	0			0							
T		0	0		0	1	0	0	0	0			0							